

ON PERMANENTS OF RANDOM MATRICES WITH POSITIVE ELEMENTS

TONĆI ANTUNOVIĆ

ABSTRACT. We study the asymptotic behavior of permanents of $n \times n$ random matrices A with independent identically distributed positive entries and prove a strong law of large numbers for $\log \text{per } A$. We calculate the values of the limit $\lim_{n \rightarrow \infty} \frac{\log \text{per } A}{n \log n}$ under the assumption that elements have power law decaying tails, and observe a first order phase transition in the limit as the mean becomes infinite. The methods extend to a wide class of rectangular matrices. It is also shown that in finite mean regime the limiting behavior holds uniformly over all submatrices of linear size.

1. INTRODUCTION

The permanent of an $m \times n$ matrix A (height m and width n) satisfying $m \leq n$ is defined as

$$\text{per } A = \sum_{\pi \in S_{m,n}} \prod_{i=1}^m a_{i,\pi(i)},$$

where $S_{m,n}$ is the set of one-to-one functions from $[m] = \{1, \dots, m\}$ to $[n] = \{1, \dots, n\}$. When $m = n$, that is when A is a square matrix, $S_{m,n} = S_n$ the set of permutations on $[n]$.

In this paper we will study asymptotics of permanents of large matrices with positive, independent and identically distributed elements. Permanents of random matrices of similar type have been studied in a number of papers.

In [2] and [3] Girko proved that

$$\lim_{n \rightarrow \infty} \frac{\log |\text{per } A| - \mathbb{E} \log |\text{per } A|}{n} \rightarrow 0,$$

(without estimating $\mathbb{E} \log |\text{per } A|$) for $n \times n$ square matrices A with independent elements, when either characteristic functions or Laplace transforms of elements is of the form $\exp(-c|t|^\alpha)$ (and some other finite man cases). Working in the context of perfect matchings on random bipartite graphs, Janson [6] proved central limit theorems for permanents of matrices with 0-1 iid elements. In a series of papers Rempala and Wołowski studied the permanents of large rectangular matrices with identically distributed elements of non-zero mean and finite variance (allowing some correlation among elements in each column). In the case of iid elements, relying on earlier results of van Es and Helmers [1] and Borovskikh and Korolyuk [7], they proved central limit theorems [10] for $(\text{per } A)/\mathbb{E} \text{per } A$ and later certain strong laws of large numbers [11]. See also Chapter 3 in [12] for a self-contained discussion of these results. Recently Tao and Vu [13] obtained significantly different behavior for $n \times n$ matrices A with independent mean zero Bernoulli ± 1 elements. They showed that with high probability $|\text{per}(A)| = n^{(\frac{1}{2} + o(1))n}$.

2010 *Mathematics Subject Classification.* Primary 60B20, 15B52, 15A15.

Key words and phrases. permanent, random matrices, law of large numbers, heavy tail distribution.

The above results demonstrate the contrast between the non-zero mean, finite variance case and the Bernoulli case, which can be summarized as

$$(1) \quad \lim_{m,n \rightarrow \infty} \frac{\log |\text{per } A|}{m \log n} = \begin{cases} 1, & \text{in the case of the finite variance and non zero mean,} \\ \frac{1}{2}, & \text{in the Bernoulli case with zero mean for } m = n. \end{cases}$$

In the non-zero finite mean case (μ being mean of the elements) the value of the limit, and especially the upper bounds, can be inferred by calculating the first moment $\mathbb{E}(\text{per } A) = \binom{n}{m} m! \mu^n$, and in the Bernoulli case from the second moment $\mathbb{E}(|\text{per } A|^2) = n!$. In this paper we will calculate the value of this limit under the assumption that elements are positive and have power law decaying tails $\mathbb{P}(\xi \geq t) = t^{-1/\beta+o(1)}$. In the case $\beta > 1$ elements have infinite mean which prevents us from guessing the value of the limit. Actually in Theorem 1 we will observe a first order phase transition in the limit at $\beta = 1$, when the mean becomes infinite.

2. SETUP AND THE RESULT

In the text we will assume that for $m \leq n$, $A_{m,n}$ is an $m \times n$ matrix (A_n when $m = n$) with independent positive elements distributed as ξ . Note that we will drop the subscripts when there is no confusion. Assuming that the matrices are constructed on a common probability space, theorems below give strong laws of large numbers for $\frac{\log \text{per } A_n}{n \log n}$ (in particular they imply a weak law of large numbers without the assumption that A_n are given on a common probability space). Extensions to rectangular matrices are given in Section 5.

Theorem 1. *Let ξ be a positive random variable satisfying*

$$(2) \quad \lim_{t \rightarrow \infty} \frac{\log \mathbb{P}(\xi \geq t)}{\log t} = -\frac{1}{\beta},$$

for some $\beta > 0$. If $(A_n)_n$ is a sequence $n \times n$ matrices on a common probability space with elements which are independent and identically distributed as ξ , then almost surely

$$(3) \quad \lim_{n \rightarrow \infty} \frac{\log \text{per } A_n}{n \log n} = \max(1, \beta).$$

Random variable ξ in (2) has finite variance for $\beta < 1/2$, finite mean for $\beta < 1$, and infinite mean for $\beta > 1$ when we observe a limit different from the values in (1).

The following result generalizes the case $\beta < 1$. It does not require the finite variance assumption and gives the general lower bounds and the upper bounds in the case of finite mean uniformly over all submatrices of linear size. Note that for an $m \times n$ matrix $A = (a_{ij})$ any matrix $B = (a_{ij})_{i \in I, j \in J}$, where $I \subset [m]$, $J \subset [n]$ is called a submatrix of A .

Theorem 2. *Assume that $(A_n)_n$ is a sequence of $n \times n$ matrices on a common probability space with elements which are independent and identically distributed as ξ , and let $0 < \alpha < 1$.*

i) *We have*

$$(4) \quad \liminf_{n \rightarrow \infty} \min_{(k,B)} \frac{\log \text{per } B}{k \log k} \geq 1,$$

where the minimum is taken over all integers $\alpha n \leq k \leq n$ and all $k \times k$ submatrices B of A_n .

ii) If ξ has a finite mean then

$$(5) \quad \limsup_{n \rightarrow \infty} \max_{(k, B)} \frac{\log \text{per } B}{k \log k} = 1,$$

where the maximum is again taken over all integers $\alpha n \leq k \leq n$ and all $k \times k$ submatrices B of A_n .

Condition (2) in Theorem 1 is satisfied with $\beta > 1$ for many common heavy tail distributions including Pareto distribution, Lévy distribution, Inverse-Gamma distribution, Beta-prime distribution and many more. We are particularly interested in the case when ξ has Pareto distribution with parameter β , that is $\mathbb{P}(\xi \geq t) = t^{-1/\beta}$ for $t \geq 1$. Actually in Section 3 the upper bounds in Theorem 1 will be proven in the Pareto case and then extended to the general case via simple stochastic domination. Note that when the convergence (2) fails to hold, one cannot guarantee the existence of the limit in (3) (see Example 16 in Section 5). However, the upper bound on the limsup in (2) will imply the upper bound in (3) and similarly the lower bound for lim inf.

Remark 3. From a more combinatorial point of view permanents can be interpreted in the context of saturated matchings (or perfect matchings for $m = n$) of bipartite graphs. For a bipartite graph $G = (V, E)$ let $V = V_1 \cup V_2$, $|V_1| \leq |V_2|$ be a decomposition of the vertex sets into subsets so that no two vertices in V_i are connected by an edge. Saturated matchings of G can be defined as subsets $\mathcal{M} \subset E$ of the edge set with the property that every vertex is adjacent to at most one edge in \mathcal{M} and that every vertex in the smaller component V_1 is adjacent to at least one edge in \mathcal{M} . For $m \leq n$ and an $m \times n$ matrix $A = (a_{ij})$ containing only elements 0 and 1, construct a bipartite graph G with $m + n$ vertices $\{v_1, \dots, v_m, w_1, \dots, w_n\}$ so that v_i and w_j are connected by an edge if and only if $a_{ij} = 1$. Clearly every one-to-one function $\pi: [m] \rightarrow [n]$ for which $\prod_{i=1}^m a_{i\pi(i)}$ is non-vanishing corresponds to a saturated matching on G in 1–1 manner. Therefore $\text{per } A$ is equal to the number of saturated matchings on G . For general matrices A one can construct the graph by drawing an edge between v_i and w_j whenever $a_{ij} \neq 0$ and putting the weight a_{ij} on this edge. Then $\text{per } A$ can be interpreted as the total weight of all the saturated matchings on G (a weight of a matching being the product of the weight on its edges). All the results in this paper can be interpreted in this way.

In the following section we prove the upper bounds in Theorem 1 and in Section 4 we provide the lower bounds and prove Theorem 2. In the last section we will extend the results to a large class of rectangular matrices and show an example demonstrating that, in general without (2) Theorem 1 fails to hold.

3. PROOF OF THE UPPER BOUNDS IN THEOREM 1

The exact calculations needed for the proof of the upper bounds in Theorem 1 are easier to perform when we are given a concrete distribution of ξ to work with. The proof will be provided for the Pareto case, but first we will see how this yields the upper bounds in Theorem 1 for the general case.

Remark 4. Throughout the paper we will use the following two simple observations.

i) For any $m \times n$ matrix A and $\lambda \in \mathbb{R}$ we have that $\text{per}(\lambda A) = \lambda^m \text{per } A$. Thus the value of the limit of $\log \text{per } A_n / (n \log n)$ in Theorem 1 (as well as lim inf and lim sup) is unchanged if we replace the generic random variable ξ by random variables $\lambda \xi$, for any $\lambda > 0$.

ii) Assume that we are given two random variables ξ_1 and ξ_2 with right continuous cumulative distribution functions $F_1(t)$ and $F_2(t)$ (and denote $F_i^-(t) = \lim_{s \uparrow t} F_i(s)$). If ξ_2 stochastically dominates ξ_1 , that is $F_2(t) \leq F_1(t)$ for any $t > 0$ (equivalently $\mathbb{P}(\xi_1 \geq t) \leq \mathbb{P}(\xi_2 \geq t)$), then one can enlarge the probability space Ω that supports ξ_1 and construct a version of ξ_2 on the larger probability space which dominates ξ_1 pointwise. For example, if ξ_1 is defined on (Ω, \mathbf{P}) then on $(\Omega \times [0, 1], \mathbf{P} \times d\lambda)$, where $d\lambda$ is the Lebesgue measure on $[0, 1]$, we define $u = \lambda F_1(\xi_1) + (1 - \lambda)F_1^-(\xi_1)$. It is easy to check that u is uniformly distributed on $[0, 1]$. Knowing this it is also easy to check that $\bar{\xi}_2 = \inf\{t : F_2(t) \geq u\}$ has the same distribution as ξ_2 and that $F_1^-(\xi_1) \leq u \leq F_2(\bar{\xi}_2)$. The last inequalities imply that $\xi_1 \leq \bar{\xi}_2$ everywhere. Thus if $(A_{m,n}^1)$ is a sequence of $m \times n$ matrices with independent elements distributed as ξ_1 defined on a common probability space, one can enlarge the probability space and construct on it a sequence of $m \times n$ matrices $(A_{m,n}^2)$ whose elements are independent and distributed as ξ_2 such that for every n , the ij th element of $A_{m,n}^1$ is not larger than the corresponding element of $A_{m,n}^2$. In particular, if ξ_1 and ξ_2 are almost surely positive then $\text{per } A_{m,n}^1 \leq \text{per } A_{m,n}^2$. The analogous claim holds when ξ_1 dominates ξ_2 .

Proof of the upper bounds in Theorem 1 assuming it holds for the Pareto case. Fix $\epsilon > 0$ and take $M > 1$ so that $\mathbb{P}(\xi \geq t) \leq t^{-1/(\beta+\epsilon)}$ holds for all $t \geq M$. Denote by $\bar{\xi}_{\beta+\epsilon}$ a Pareto distributed random variable with parameter $\beta + \epsilon$ and observe that $\mathbb{P}(\xi \geq t) \leq \mathbb{P}(M\bar{\xi}_{\beta+\epsilon} \geq t)$ holds for all t . Assuming the statement holds for the Pareto case, Remark 4 implies that almost surely

$$\limsup_n \frac{\log \text{per } A_n}{n \log n} \leq \beta + \epsilon.$$

Since $\epsilon > 0$ was arbitrary the claim follows. \square

The rest of this section is devoted to the proof of the upper bounds in the Pareto case in which we show explicit calculations. A useful observation which we will use extensively is the fact that if ξ is a Pareto distributed random variable with parameter β , then $Y = (\log \xi)/\beta$ has exponential distribution with rate 1, that is $\mathbb{P}(Y \geq t) = e^{-t}$ for $t \geq 0$. We start by proving some basic estimates for maxima of independent exponential random variables.

Lemma 5. *Let $n \geq 2$ and Y_i , $1 \leq i \leq n$ be independent exponential random variables with rate 1 and $R = \frac{\max_{1 \leq i \leq n} Y_i}{\log n}$.*

i) *For any positive t*

$$(6) \quad \mathbb{P}(R \leq t) \leq \exp(-n^{1-t}), \text{ and } \mathbb{P}(R \geq t) \leq n^{1-t}.$$

ii) *The expectation of e^R can be bounded as*

$$(7) \quad \mathbb{E}(e^R) \leq \exp\left(1 + \frac{1}{\log n - 1}\right).$$

Proof. i) Both inequalities are straightforward. First we calculate

$$(8) \quad \mathbb{P}(R \leq t) = \mathbb{P}\left(\max_{1 \leq i \leq n} Y_i \leq t \log n\right) = \left(1 - e^{-t \log n}\right)^n = (1 - n^{-t})^n.$$

Now applying the inequality $1 - x \leq e^{-x}$ to the right hand side of (8) we obtain the first inequality in (6). Using (8) and the inequality $(1 - x)^n \geq 1 - nx$ which holds for all positive integers n and all $0 < x < 1$ we prove the second inequality in (6):

$$\mathbb{P}(R \geq t) = 1 - (1 - n^{-t})^n \leq n^{1-t}.$$

ii) Using the second inequality in (6) we obtain for $t \geq 1$

$$\mathbb{P}(e^R \geq t) = \mathbb{P}(R \geq \log t) \leq \frac{n}{n^{\log t}} = \frac{n}{t^{\log n}},$$

from where we get

$$\begin{aligned} \mathbb{E}(e^R) &= \int_0^\infty \mathbb{P}(e^R \geq t) dt \leq e + \int_e^\infty \frac{n}{t^{\log n}} dt = e + \frac{n}{\log n - 1} \frac{1}{e^{\log n - 1}} \\ &= e \left(1 + \frac{1}{\log n - 1} \right) \leq \exp \left(1 + \frac{1}{\log n - 1} \right) \end{aligned}$$

□

The idea of the proof of the upper bounds in the Pareto case is to estimate (by evaluating the expectation) the number of permutations π for which the product $\prod_{i=1}^n \xi_{i\pi(i)}$ will lie in some given interval. The key estimate is provided in Lemma 7. We will only consider the intervals not exceeding $(n\sqrt{\log n})^{\beta n}$, since as the following lemma shows, the largest product $\prod_i \xi_{i\pi(i)}$ typically does not exceed this value.

Lemma 6. *If $(Y_{i,j})_{i,j}$ are independent exponential random variables with rate 1 then for any $\lambda > 0$*

$$(9) \quad \sum_{n=2}^\infty \mathbb{P} \left(\max_{\pi \in S_n} \sum_{i=1}^n Y_{i,\pi(i)} \geq n \log n + n \frac{\log \log n}{\lambda} \right) < \infty.$$

Proof. Denote $R_i := \frac{\max_{1 \leq j \leq n} Y_{i,j}}{\log n}$. From the definition of R_i it is obvious that

$$\max_{\pi \in S_n} \sum_{i=1}^n Y_{i,\pi(i)} \leq \log n \left(\sum_{i=1}^n R_i \right).$$

Using the inequality (7) we have

$$\begin{aligned} \mathbb{P} \left(\max_{\pi \in S_n} \sum_{i=1}^n Y_{i,\pi(i)} \geq n \log n + n \frac{\log \log n}{\lambda} \right) &\leq \mathbb{P} \left(\sum_{i=1}^n R_i \geq n + n \frac{\log \log n}{\lambda \log n} \right) \\ &\leq \mathbb{E} \left(e^{\sum_{i=1}^n R_i - n - n \frac{\log \log n}{\lambda \log n}} \right) = \left(\frac{\mathbb{E}(e^{R_1})}{e^{(\log n) \frac{1}{\lambda \log n}}} \right)^n \leq \exp \left(\left(\frac{1}{\log n - 1} - \frac{\log \log n}{\lambda \log n} \right) n \right). \end{aligned}$$

The right hand side above is summable in n which proves the lemma. □

Lemma 7. *Let $(Y_{i,j})_{i,j}$ be independent exponential random variables with rate 1 and*

$$(10) \quad Z_{n,k} = \left| \left\{ \pi \in S_n : (k-1)n \leq \sum_{i=1}^n Y_{i,\pi(i)} < kn \right\} \right|.$$

Then for any $\gamma > 1$ and $\lambda > 1$ we have

$$(11) \quad \sum_{n=2}^\infty \mathbb{P} \left(\bigcup_{1 \leq k \leq \log n + \frac{\log \log n}{\lambda}} \{Z_{n,k} > \mathbb{E}(Z_{n,k})^\gamma\} \right) < \infty.$$

Proof. First by Markov's inequality

$$\begin{aligned} \mathbb{P}\left(\bigcup_{1 \leq k \leq \log n + \frac{\log \log n}{\lambda}} \{Z_{n,k} > \mathbb{E}(Z_{n,k})^\gamma\}\right) &\leq \sum_{1 \leq k \leq \log n + \frac{\log \log n}{\lambda}} \mathbb{P}(Z_{n,k} > \mathbb{E}(Z_{n,k})^\gamma) \\ &\leq \sum_{1 \leq k \leq \log n + \frac{\log \log n}{\lambda}} \frac{1}{\mathbb{E}(Z_{n,k})^{\gamma-1}}. \end{aligned}$$

Let's calculate the expectation of $Z_{n,k}$. Clearly for any fixed $\pi \in S_n$ we have

$$\mathbb{E}(Z_{n,k}) = n! \mathbb{P}\left((k-1)n \leq \sum_{i=1}^n Y_{i,\pi(i)} < kn\right).$$

Since for a fixed π , $\sum_{i=1}^n Y_{i,\pi(i)}$ is the sum of n independent exponential random variables with mean 1, it has Gamma density $\frac{x^{n-1}e^{-x}}{(n-1)!}$. Therefore the expectation of $Z_{n,k}$ is given by

$$(12) \quad \mathbb{E}(Z_{n,k}) = n! \int_{(k-1)n}^{kn} \frac{x^{n-1}e^{-x}}{(n-1)!} dx = n \int_{(k-1)n}^{kn} x^{n-1}e^{-x} dx.$$

In particular we have

$$(13) \quad \mathbb{E}(Z_{n,k}) \leq e^{-(k-1)n} \int_{(k-1)n}^{kn} nx^{n-1} dx \leq (kn)^n e^{-(k-1)n}.$$

Furthermore by (12)

$$(14) \quad \mathbb{E}(Z_{n,k}) = n \int_{(k-1)n}^{kn} x^{n-1}e^{-x} dx \geq e^{-kn} \int_{(k-1)n}^{kn} nx^{n-1} dx = e^{-kn} n^n (k^n - (k-1)^n).$$

In particular $\mathbb{E}(Z_{n,1}) \geq \left(\frac{n}{e}\right)^n$, which implies that the series $\sum_{n=1}^{\infty} \mathbb{E}(Z_{n,1})^{1-\gamma}$ converges to a finite limit. Therefore we are left to prove

$$(15) \quad \sum_{n=2}^{\infty} \sum_{2 \leq k \leq \log n + \frac{\log \log n}{\lambda}} \frac{1}{\mathbb{E}(Z_{n,k})^{\gamma-1}} < \infty.$$

Again using (14) and the inequality $k^n \geq (k-1)^n + n(k-1)^{n-1}$ we obtain

$$\mathbb{E}(Z_{n,k}) \geq e^{-kn} n^{n+1} (k-1)^{n-1} \geq e^{-kn} n^n (k-1)^n,$$

from where

$$(16) \quad \sum_{2 \leq k \leq \log n + \frac{\log \log n}{\lambda}} \frac{1}{\mathbb{E}(Z_{n,k})^{\gamma-1}} \leq \frac{e^{n(\gamma-1)}}{n^{n(\gamma-1)}} \sum_{1 \leq k \leq \log n + \frac{\log \log n}{\lambda}} \frac{e^{kn(\gamma-1)}}{k^{n(\gamma-1)}}.$$

The function $g(t) = e^{tn(\gamma-1)} t^{-n(\gamma-1)}$ is convex since

$$g''(t) = n(\gamma-1) \frac{e^{tn(\gamma-1)}}{t^{n(\gamma-1)+2}} (n(\gamma-1)(t-1)^2 + 1) \geq 0.$$

Therefore for any $1 \leq k \leq \log n + \frac{\log \log n}{\lambda}$ we have

$$(17) \quad \frac{e^{kn(\gamma-1)}}{k^{n(\gamma-1)}} = g(k) \leq \max \left\{ g(1), g\left(\log n + \frac{\log \log n}{\lambda}\right) \right\}.$$

For any n large enough we have

$$g\left(\log n + \frac{\log \log n}{\lambda}\right) = \frac{n^{n(\gamma-1)}(\log n)^{\frac{n(\gamma-1)}{\lambda}}}{(\log n + \frac{\log \log n}{\lambda})^{n(\gamma-1)}} \geq \left(\frac{n}{2(\log n)^{1-1/\lambda}}\right)^{n(\gamma-1)} \geq e^{n(\gamma-1)} = g(1),$$

and, for such n , using (17) we also get

$$\frac{e^{kn(\gamma-1)}}{k^{n(\gamma-1)}} \leq g\left(\log n + \frac{\log \log n}{\lambda}\right) = \frac{n^{n(\gamma-1)}(\log n)^{\frac{n(\gamma-1)}{\lambda}}}{(\log n + \frac{\log \log n}{\lambda})^{n(\gamma-1)}} \leq \frac{n^{n(\gamma-1)}}{(\log n)^{n(\gamma-1)(1-1/\lambda)}}.$$

Thus, for n large enough (16) yields

$$\begin{aligned} \sum_{2 \leq k \leq \log n + \frac{\log \log n}{\lambda}} \frac{1}{\mathbb{E}(Z_{n,k})^{\gamma-1}} &\leq \frac{e^{n(\gamma-1)}}{n^{n(\gamma-1)}} \frac{n^{n(\gamma-1)}}{(\log n)^{n(\gamma-1)(1-1/\lambda)}} \left(\log n + \frac{\log \log n}{\lambda}\right) \\ &= \left(\frac{e}{(\log n)^{1-1/\lambda}}\right)^{n(\gamma-1)} \left(\log n + \frac{\log \log n}{\lambda}\right). \end{aligned}$$

The expression on the right hand side is summable in n which proves (15) and thus also (11). \square

Now we are ready to finish the proof of upper bounds.

Proof of the upper bounds in Theorem 1 for the Pareto case. Define $Z_{n,k}$ as in (10) and fix an arbitrary $\gamma > 1$. Lemmas 6 and 7 and Borel-Cantelli lemma imply that almost surely there exists a positive integer n_0 such that for all $n \geq n_0$ we have

$$\begin{aligned} \max_{\pi \in S_n} \sum_{i=1}^n Y_{i,\pi(i)} &\leq n \log n + \frac{n \log \log n}{2} \quad \text{and} \\ Z_{n,k} &\leq \mathbb{E}(Z_{n,k})^\gamma, \quad \text{for each } 1 \leq k \leq \log n + \frac{\log \log n}{2}. \end{aligned}$$

Using (13) the following inequalities are almost surely satisfied for n large enough

$$\begin{aligned} (18) \quad \text{per } A &= \sum_{\pi \in S_n} e^{\beta \sum_{i=1}^n Y_{i,\pi(i)}} \leq \sum_{1 \leq k \leq \log n + \frac{\log \log n}{2}} e^{\beta kn} Z_{n,k} \leq \sum_{1 \leq k \leq \log n + \frac{\log \log n}{2}} e^{\beta kn} \mathbb{E}(Z_{n,k})^\gamma \\ &\leq e^{\gamma n} n^{\gamma n} \sum_{1 \leq k \leq \log n + \frac{\log \log n}{2}} k^{\gamma n} e^{(\beta-\gamma)kn} \\ &\leq e^{\gamma n} n^{\gamma n} \left(\log n + \frac{\log \log n}{2}\right) \max_{1 \leq \tau \leq \log n + \frac{\log \log n}{2}} \left(\tau^{\gamma n} e^{(\beta-\gamma)\tau n}\right). \end{aligned}$$

If $\beta > 1$ and γ is such that $\beta > \gamma > 1$, $\tau^{\gamma n} e^{(\beta-\gamma)\tau n}$ is an increasing function in τ and thus for n large enough

$$\text{per } A \leq e^{\gamma n} n^{\beta n} (\log n)^{\frac{(\beta-\gamma)n}{2}} \left(\log n + \frac{\log \log n}{2}\right)^{\gamma n+1}.$$

This yields

$$\frac{\log \text{per } A}{n \log n} \leq \frac{\gamma}{\log n} + \beta + \frac{\beta - \gamma}{2} \frac{\log \log n}{\log n} + \left(\gamma + \frac{1}{n}\right) \frac{\log \left(\log n + \frac{\log \log n}{2}\right)}{\log n},$$

from where clearly

$$\limsup_{n \rightarrow \infty} \frac{\log \text{per } A}{n \log n} \leq \beta.$$

In the case $\beta \leq 1$ we want to maximize the function $\tau^{\gamma n} e^{(\beta - \gamma)\tau n}$. Write $e^{h(\tau)} := \tau^{\gamma n} e^{(\beta - \gamma)\tau n}$. We get

$$h(\tau) = \gamma n \log \tau + (\beta - \gamma)\tau n, \quad h'(\tau) = \frac{\gamma n}{\tau} + (\beta - \gamma)n, \quad h''(\tau) = -\frac{\gamma n}{\tau^2} < 0.$$

Therefore function h is concave and the maximum occurs when $h'(\tau) = 0$, that is when $\tau = \frac{\gamma}{\gamma - \beta}$ at which the value of the function $e^{h(\tau)}$ is equal to $\left(\frac{\gamma}{(\gamma - \beta)e}\right)^{\gamma n}$. From (18) we get

$$\text{per } A \leq n^{\gamma n} \left(\log n + \frac{\log \log n}{2} \right) \left(\frac{\gamma}{\gamma - \beta} \right)^{\gamma n}$$

and so

$$\frac{\log \text{per } A}{n \log n} \leq \gamma + \frac{\log \left(\log n + \frac{\log \log n}{2} \right)}{n \log n} + \gamma \frac{\log \gamma - \log(\gamma - \beta)}{\log n} \rightarrow \gamma,$$

as $n \rightarrow \infty$. Since $\gamma > 1$ was arbitrary the claim follows. \square

4. LOWER BOUNDS AND THE PROOF OF THEOREM 2

In this section we prove Theorem 2 as well as the lower bounds in Theorem 1. An important ingredient is the use of stochastic domination to reduce certain technical issues to iid 0, 1 matrices. The following result proven by Hall [4] and Mann and Ryser in [8] provides lower bounds for permanents of such matrices (see also Theorem 1.2 in Chapter 4 of [9]).

Proposition 8. *Let A be an $m \times n$ matrix, $m \leq n$ whose all elements are equal to 0 or 1. Assume that each row of A contains at least k elements equal to 1. If $k \geq m$, then*

$$(19) \quad \text{per } A \geq \frac{k!}{(k - m)!}.$$

If $k < m$ and $\text{per } A > 0$ then

$$(20) \quad \text{per } A \geq k!.$$

As discussed in the introduction (see Remark 3) permanents of matrices with 0, 1 elements can be viewed as the number of saturated matchings on corresponding bipartite graphs. To ensure the positivity of the permanent, when applying (20), we will exploit this connection through the classical Hall's marriage theorem, which can be easily stated in this setting (see [5]).

Theorem 9. *Let $G = (V, E)$ be a bipartite graph and let $V = V_1 \cup V_2$ be a decomposition of the vertex set so that no two vertices in V_i are connected by an edge, $i = 1, 2$. Assuming $|V_1| \leq |V_2|$, there exists a saturated matching on G if and only if for any subset $W \subset V_1$ we have $|W| \leq |\{v : v \sim w, w \in W\}|$.*

Restating the above theorem in terms of permanents of 0, 1 matrices yields the following lemma.

Lemma 10. *Let B be an $m \times n$ matrix whose all elements are either 0 or 1. If for any $1 \leq k \leq m$ any $k \times (n - k + 1)$ submatrix of B has at least one element equal to 1, then $\text{per } B \geq 1$.*

All the necessary applications of Proposition 8 and Lemma 10 are summarized in Lemma 12 which, in particular, proves the lower bounds in Theorem 2.

Remark 11. Recall that Stirling's formula says that

$$\lim_{n \rightarrow \infty} n! e^n n^{-(n+1/2)} = \sqrt{2\pi}.$$

In particular there are constants $c_1 < c_2$ so that for any n and $1 \leq k \leq n-1$

$$(21) \quad c_1 \frac{n^{n+1/2}}{k^{k+1/2}(n-k)^{n-k+1/2}} \leq \binom{n}{k} \leq c_2 \frac{n^{n+1/2}}{k^{k+1/2}(n-k)^{n-k+1/2}}.$$

Lemma 12. *Let ξ be a positive random variable and let A_n be a sequence of $n \times n$ matrices whose elements are independent and identically distributed as ξ . For any $0 < \alpha < 1$ and any $\delta > 0$ there exists $r > 0$ with the following property: Almost surely there exists n_0 such that for any $n \geq n_0$ and any $\alpha n \leq k \leq n$, any $k \times k$ submatrix B of A_n satisfies $\text{per } B \geq r^k k^{(1-\delta)k}$.*

Proof. Let $q > 0$ be such that $\mathbb{P}(\xi \leq q) < \eta$, where η is to be chosen later. Define the random variable $\tilde{\xi} = \mathbf{1}_{(\xi \geq q)}$. and define the matrix $\tilde{A}_n = (\tilde{\xi}_{ij})$. Let \mathfrak{B}_n denote the event that some row of \tilde{A}_n contains more than $\alpha \delta n$ zeros and let \mathfrak{C}_n denote the event that for some k_1 and k_2 satisfying $\alpha n \leq k_1 + k_2$ there exists a $k_1 \times k_2$ submatrix of \tilde{A}_n containing only zeros. By Lemma 10 on the event \mathfrak{C}_n^c any $k \times k$ submatrix of \tilde{A}_n has a positive permanent, for $\alpha n \leq k \leq n$. Furthermore on the event \mathfrak{B}_n^c every $k \times k$ submatrix of \tilde{A}_n for $k \geq \alpha n$ contains at least $(1-\delta)k$ ones. Thus on the event $\mathfrak{B}_n^c \cap \mathfrak{C}_n^c$ by (20) we have for any $k \geq \alpha n$ and any $k \times k$ submatrix B of A_n

$$\text{per } B \geq q^k \text{per } \tilde{B} \geq q^k \lfloor (1-\delta)k \rfloor! \geq \left(\frac{q(1-\delta)}{e} \right)^k k^{(1-\delta)k},$$

where \tilde{B} is the submatrix of \tilde{A}_n having the same rows and columns as B in A_n . Note that the last inequality above holds for n large enough by Stirling's approximation.

Thus we only need to prove that the probabilities of the events $\mathfrak{B}_n \cup \mathfrak{C}_n$ are summable (since then they happen only finitely many times almost surely). To end this observe that the average number of 1s in every row and column of \tilde{A}_n is greater than $n(1-\eta)$, so for $\eta < \alpha\delta$ by standard large deviation arguments there exists a constant C such that $\mathbb{P}(\mathfrak{B}_n) \leq Cne^{-n(\alpha\delta-\eta)/C}$, which is clearly summable. For \mathfrak{C}_n use union bound to obtain

$$(22) \quad \begin{aligned} \mathbb{P}(\mathfrak{C}_n) &\leq 2 \sum_{\substack{\alpha n \leq k_1 + k_2 \leq n \\ k_1 \geq k_2}} \binom{n}{k_1} \binom{n}{k_2} \eta^{k_1 k_2} \leq 2 \sum_{\alpha n/2 \leq k_1 \leq n} \binom{n}{k_1} \sum_{1 \leq k_2 \leq n} \binom{n}{k_2} \eta^{k_1 k_2} \\ &\leq 2 \sum_{\alpha n/2 \leq k_1 \leq n} \binom{n}{k_1} \left((1 + \eta^{k_1})^n - 1 \right) \leq 2 \sum_{\alpha n/2 \leq k_1 \leq n} \binom{n}{k_1} (2\eta)^{k_1}, \end{aligned}$$

for η small enough. It is easy to check the last inequality by writing $(1 + \eta^{k_1})^n - 1 = \eta^{k_1} \sum_{\ell=0}^{n-1} (1 + \eta^{k_1})^\ell$ and bounding each of the terms on the right hand side. To prove that the right hand side above is summable, observe that for $\eta = \eta(\alpha)$ sufficiently small the following inequalities hold for $\alpha n/2 \leq k_1 \leq n$

$$(2\eta)^{k_1/2} \leq (2\eta)^{\alpha n/4} \leq (1 - \sqrt{2\eta})^{(1-\alpha/2)n} \leq (1 - \sqrt{2\eta})^{n-k_1}.$$

Plugging this back into (22) we get

$$\mathbb{P}(\mathfrak{C}_n) \leq 2 \sum_{\alpha n/2 \leq k_1 \leq n} \binom{n}{k_1} (2\eta)^{k_1/2} (1 - \sqrt{2\eta})^{n-k_1}.$$

The right hand side is just twice the probability that the sum of n independent random variables having value 1 with probability $\sqrt{2\eta}$ and value 0 otherwise, is greater than $\alpha n/2$. Choosing $\sqrt{2\eta} < \alpha/2$, large deviation principle implies that this probability is exponentially small, and thus summable in n . This finishes the proof. \square

Proof of Theorem 2. i) Taking an arbitrary $\delta > 0$ by Lemma 12 we can find $r > 0$ small enough so that almost surely for n large enough

$$\frac{\log \text{per } B}{k \log k} \geq \frac{\log r}{\log k} + (1 - \delta),$$

for any $\alpha n \leq k \leq n$ and any $k \times k$ submatrix B of A_n . Thus almost surely

$$\liminf_n \min_{B,k} \frac{\log \text{per } B}{k \log k} \geq 1 - \delta.$$

Since $\delta > 0$ was arbitrary the claim follows.

ii) Let $\tilde{\xi}$ be parameter 1 Pareto distributed random variable. By Markov inequality, for all $t \geq \mathbb{E}(\xi)$ we have

$$\mathbb{P}(\xi \geq t) \leq \frac{\mathbb{E}(\xi)}{t} = \mathbb{P}(\mathbb{E}(\xi)\tilde{\xi} \geq t),$$

Thus $\mathbb{E}(\xi)\tilde{\xi}$ stochastically dominates ξ from above and by Remark 4 we can construct a sequence (\tilde{A}_n) of random $n \times n$ matrices whose elements are independent and identically distributed as $\tilde{\xi}$, such that for any $\alpha n \leq k \leq n$ and any $k \times k$ submatrix B of A_n for the corresponding submatrix \tilde{B} of \tilde{A}_n we have $\text{per } B \leq \mathbb{E}(\xi)^k \text{per } \tilde{B}$. Thus it is enough to prove the claim in the case when we replace ξ with $\tilde{\xi}$. If \tilde{B} is an arbitrary $k \times k$ submatrix of \tilde{A}_n then after permuting the rows and columns we can assume that \tilde{B} is at the intersection of the first k rows and columns. Denote by \tilde{B}^c the matrix at the intersection of the other $n - k$ rows and columns and observe that since \tilde{A}_n has positive elements $\text{per } \tilde{A}_n \geq \text{per } \tilde{B} \text{per } \tilde{B}^c$. Furthermore since all the elements are larger than 1 we have $\text{per } \tilde{B}^c \geq (n - k)!$ and thus

$$\frac{\text{per } \tilde{B}}{k \log k} \leq \frac{\text{per } \tilde{A}_n}{k \log k} - \frac{\log(n - k)!}{k \log k} \leq \frac{\text{per } \tilde{A}_n}{k \log k} - \frac{(n - k)(\log(n - k) - 1) - c}{k \log k},$$

for some $c > 0$, where the second inequality follows from Stirling's formula (note that we can assume that $k < n$, since for $k = n$ the upper bounds have been proven in the previous section). By the upper bounds in Theorem 1, for any $\epsilon > 0$ almost surely there is n_0 such that for all $n \geq n_0$ we have $\text{per } \tilde{A}_n \leq (1 + \epsilon)n \log n$. For such n we have

$$\begin{aligned} \frac{\text{per } \tilde{B}}{k \log k} &\leq (1 + \epsilon) \frac{n \log n}{k \log k} - \frac{(n - k) \log(n - k)}{k \log k} + \frac{n - k}{k \log k} + \frac{c}{k \log k} \\ &\leq \frac{1}{\alpha} \left(\frac{(1 + \epsilon) \log n}{\log n + \log \alpha} - 1 \right) + \frac{n - k}{k \log k} + \frac{c}{k \log k} + \frac{n}{k} - \left(\frac{n}{k} - 1 \right) \frac{\log(n - k)}{\log k}. \end{aligned}$$

For a fixed α the first three terms on the right hand side vanish in the limit and it suffices to show that

$$(23) \quad \limsup_{n \rightarrow \infty} \max_{2 \leq k < n} \left(\frac{n}{k} - \left(\frac{n}{k} - 1 \right) \frac{\log(n-k)}{\log k} \right) \leq 1.$$

Denoting $n = tk$, where $t \geq 1$ we have

$$\frac{n}{k} - \left(\frac{n}{k} - 1 \right) \frac{\log(n-k)}{\log k} = 1 - \frac{(t-1) \log(t-1)}{\log k}.$$

When taking supremum one can assume that $1 < t < 2$, that is $k > n/2$. In that case the claim follows from the fact that $s \mapsto s \log s$ is bounded on $(0, 1)$. \square

Proof of the lower bounds in Theorem 1. The lower bounds for $\beta \leq 1$ follow from Theorem 2 i), so in the rest of the proof we will assume that $\beta > 1$.

First define the random variable $Y = (\log \xi)/\beta$ and observe that

$$\lim_{t \rightarrow \infty} \frac{\log \mathbb{P}(Y \geq t)}{t} = -1,$$

and thus for any $\epsilon > 0$ we have $\mathbb{P}(Y \geq t) \geq \exp(-t(1+\epsilon))$, for t large enough. Let (Y_{ij}) be an array of independent random variables distributed as Y , and for a fixed n define

$$Q_k = \max_{1 \leq i \leq n-k+1} Y_{ki}, \text{ for } 1 \leq k \leq n.$$

Following the same arguments as in the proof of the first inequality in (6) one can prove that for any $t' > 0$

$$\mathbb{P}(Q_1 \leq t' \log n) \leq \exp(-n^{1-t'(1+\epsilon)}),$$

holds for n large enough. In particular if we fix $0 < t' < 1$, $\epsilon > 0$ such that $t'(1+\epsilon) < 1$ and $0 < \rho < 1$ then for n large enough

$$(24) \quad \sum_{1 \leq k \leq \rho n} \mathbb{P} \left(\frac{Q_k}{\log(n-k+1)} \leq t' \right) \leq \sum_{1 \leq k \leq \rho n} \exp \left(- (n-k+1)^{1-t'(1+\epsilon)} \right) \\ \leq \rho n \exp \left(- (1-\rho)^{1-t'(1+\epsilon)} n^{1-t'(1+\epsilon)} \right).$$

We will need this estimate later in the proof.

Now we will run a greedy algorithm to extract a large submatrix of A_n with a large permanent. Starting from the first row of A_n we pick a largest possible admissible element in each row. First define $A^{(1)} = A_n$ and m_1 as the smallest index between 1 and n , such that $\zeta_1 := \xi_{1,m_1} = \max_{1 \leq i \leq n} \xi_{1,i}$. Next we inductively construct the matrix $A^{(k+1)}$ from $A^{(k)}$ by deleting the first row and the m_k -th column of the matrix $A^{(k)}$. At each step ζ_k is defined as the largest element in the first row of the matrix $A^{(k)}$ and m_k as the first column which contains such an element. Note that conditioned on the elements of the first $k-1$ rows of the matrix A_n the matrix $A^{(k)}$ is just an $(n-k+1) \times (n-k+1)$ matrix with independent elements distributed as ξ . This immediately implies that ζ_1, \dots, ζ_n are independent and that the first row in $A^{(k)}$ is distributed as $(\xi_{k,1}, \dots, \xi_{k,n-k+1})$. Therefore

$$(25) \quad \left(\frac{\log \zeta_1}{\beta}, \dots, \frac{\log \zeta_n}{\beta} \right) \stackrel{d}{=} (Q_1, \dots, Q_n).$$

Now fix some $0 < t < t' < 1$ and $t/t' < \rho < 1$. Then for n large enough

$$(26) \quad t \frac{n \log n}{\log n + \log(1-\rho)} \leq t'(n\rho - 1).$$

Let $k = \lfloor \rho n \rfloor$ and consider the $(n - k) \times (n - k)$ submatrix $B = A^{(k+1)}$ which is at the intersection of the last $(n - k)$ rows and columns $[n] \setminus \{m_i : 1 \leq i \leq k\}$. Consider the complement submatrix B^c which lies at the intersection of the first k rows and columns $\{m_i : 1 \leq i \leq k\}$. Since ζ_i is the largest element in the i -th row of B^c we have $\text{per } B^c \geq \prod_{i=1}^k \zeta_i$, and by (25) we have that

$$(27) \quad \mathbb{P}\left(\frac{\log \text{per } B^c}{\beta n \log n} \leq t\right) \leq \mathbb{P}\left(\sum_{i=1}^k Q_i \leq tn \log n\right) \leq \mathbb{P}\left(\sum_{i=1}^k \frac{Q_i}{\log(n - i + 1)} \leq \frac{tn \log n}{\log((1 - \rho)n)}\right).$$

Now using (26) we see that if the event under the probability on the right hand side happens, then $Q_i \leq t' \log(n - i + 1)$ happens for some $1 \leq i \leq k$. Therefore (24) and the sub-additivity imply that for any $\epsilon > 0$ such that $t'(1 + \epsilon) < 1$ we have

$$(28) \quad \mathbb{P}\left(\frac{\log \text{per } B^c}{\beta n \log n} \leq t\right) \leq \sum_{i=1}^k \mathbb{P}\left(\frac{Q_i}{\log(n - i + 1)} \leq t'\right) \leq \rho n \exp(-(1 - \rho)^{1-t'(1+\epsilon)} n^{1-t'(1+\epsilon)}),$$

for n large enough. Since the right hand side is summable in n we see that almost surely $\frac{\log \text{per } B^c}{n \log n} \leq \beta t$ happens for only finitely many integers n .

On the other hand, Lemma 12 implies that for any $\delta > 0$ there is $r > 0$ such that almost surely

$$\text{per } B \geq r^{n-k} (n - k)^{(1-\delta)(n-k)} \geq \left(r(1 - \rho)^{1-\delta}\right)^{(1-\rho)n} n^{(1-\delta)(1-\rho)n}.$$

Since A_n has positive elements, the inequality $\text{per } A_n \geq \text{per } B \text{ per } B^c$ holds and thus almost surely for n large enough

$$\frac{\log \text{per } A_n}{n \log n} \geq \beta t + (1 - \delta)(1 - \rho) + \frac{(1 - \rho)(\log r + (1 - \delta) \log(1 - \rho))}{\log n}.$$

Taking the limit as $n \rightarrow \infty$ and then $t \uparrow 1$ (which forces $\rho \uparrow 1$) yields the claim. \square

5. NON-SQUARE MATRICES AND THE NECESSITY OF (2)

In this section we sketch how the above arguments extend to a large class of rectangular matrices. We will still assume that elements are sampled independently from a distribution supported on \mathbb{R}^+ , but will now allow the width of the matrix to be significantly larger than the height, in particular it will suffice for the height to grow polynomially in the logarithm of the width. The precise condition under the method extends is that matrix A_n is $m_n \times n$, that is has height m_n and width n , and the height satisfies the condition

$$(29) \quad \liminf_n \frac{m_n \log \log n}{(\log n)^2} > 1.$$

Observe that for an $m \times n$ matrix with iid elements of mean μ we have $\mathbb{E}(\text{per } A_n) = \binom{n}{m} m! \mu^m$ which demonstrates that the scaling function $n \log n$ will have to be replaced by $m_n \log n$.

Theorem 13. *Let ξ be a positive random variable satisfying (2) for some $\beta > 0$. If $(A_n)_n$ is a sequence of $m_n \times n$ matrices on a common probability space with elements which are independent and identically distributed as ξ and satisfying (29), then almost surely*

$$(30) \quad \lim_{n \rightarrow \infty} \frac{\log \text{per } A_n}{m_n \log n} = \max(1, \beta).$$

The uniformity over all submatrices of linear size holds as well.

Theorem 14. *Assume that (A_n) is a sequence of $m_n \times n$ matrices on a common probability space with elements which are independent and identically distributed as ξ and satisfying (29), and let $0 < \alpha < 1$.*

i) *We have*

$$(31) \quad \liminf_{n \rightarrow \infty} \min_{(k_1, k_2, B)} \frac{\log \text{per } B}{k_1 \log k_2} \geq 1,$$

where the minimum is taken over all pairs of integers (k_1, k_2) satisfying $\alpha m_n \leq k_1 \leq m_n$, $\alpha n \leq k_2 \leq n$ and $k_1 \leq k_2$ and all $k_1 \times k_2$ submatrices B of A_n .

ii) *If ξ has a finite mean then*

$$(32) \quad \limsup_{n \rightarrow \infty} \max_{(k_1, k_2, B)} \frac{\log \text{per } B}{k_1 \log k_2} = 1,$$

where the minimum is taken over all pairs of integers (k_1, k_2) satisfying $\alpha m_n \leq k_1 \leq m_n$, $\alpha n \leq k_2 \leq n$ and $k_1 \leq k_2$ and all $k_1 \times k_2$ submatrices B of A_n .

The proofs of these theorems are modifications of the arguments in the previous two sections. We will briefly sketch how the modifications go and at which points one needs to do a more careful analysis. Note that, to simplify notation, we will drop the ceiling and the floor notation throughout the section.

Sketch of the proof of the upper bounds in Theorem 13. As before, by stochastic domination, it suffices to prove the claim when elements are Pareto distributed. To end this one needs to prove a version of Lemma 6 which states that when $(Y_{i,j})$ are independent exponentially distributed with rate one and $\lambda > 1$, we have

$$\sum_{n=2}^{\infty} \mathbb{P} \left(\max_{\pi \in S_{m_n, n}} \sum_{i=1}^{m_n} Y_{i, \pi(i)} \geq m_n \log n + m_n \frac{\log \log n}{\lambda} \right) < \infty.$$

Proceeding as in the proof of Lemma 6 one is left to show that

$$\sum_{n=2}^{\infty} \exp \left(\left(\frac{1}{\log n - 1} - \frac{\log \log n}{\lambda \log n} \right) m_n \right) < \infty,$$

for some $\lambda > 1$. By (29) for $\lambda > 1$ small enough the expression in the exponent is not larger than $-\alpha \log n$ for some $\alpha > 1$ which yields the claim.

Next one defines the analog of (10) as

$$(33) \quad Z_{n,k} = \left| \left\{ \pi \in S_{m_n, n} : (k-1)m_n \leq \sum_{i=1}^{m_n} Y_{i, \pi(i)} < km_n \right\} \right|,$$

and needs to prove (11). Calculating expectation of $Z_{n,k}$ as in (12) yields

$$(34) \quad \binom{n}{m_n} e^{-km_n} m_n^{m_n} (k^{m_n} - (k-1)^{m_n}) \leq \mathbb{E}(Z_{n,k}) \leq \binom{n}{m_n} e^{-(k-1)m_n} (km_n)^{m_n}.$$

For $k = 1$ this yields $\mathbb{E}(Z_{n,1}) \geq \binom{n}{m_n} m_n! e^{-m_n}$ which handles the sum $\sum_{n \geq 2} \mathbb{E}(Z_{n,1})^{1-\gamma}$. We are left to prove the analog of (16), that

$$\sum_{2 \leq k \leq \log n + \frac{\log \log n}{\lambda}} \frac{1}{\mathbb{E}(Z_{n,k})^{\gamma-1}} \leq \frac{e^{m_n(\gamma-1)}}{\binom{n}{m_n}^{\gamma-1} m_n^{m_n(\gamma-1)}} \sum_{1 \leq k \leq \log n + \frac{\log \log n}{\lambda}} \frac{e^{km_n(\gamma-1)}}{k^{m_n(\gamma-1)}}$$

is summable in n . Again by the convexity of $g(t) = e^{tm_n(\gamma-1)} t^{-m_n(\gamma-1)}$ and the fact that $g(1) \leq g(\log n + \frac{1}{\lambda} \log \log n)$, proceeding as before one is left to prove that

$$\left(\frac{ne}{m_n(\log n)^{1-1/\lambda}} \right)^{m_n(\gamma-1)} \frac{\log n + \frac{1}{\lambda} \log \log n}{\binom{n}{m_n}^{\gamma-1}}$$

is summable in n . Since $(\log n)^{-\kappa m_n}$ is summable, for any $\kappa > 0$, it is enough to show that $\binom{n}{m_n} \geq (cn/m_n)^{m_n}$, for some $c > 0$. Taking logs on both sides and using (21) (we can assume that $m_n < n$) it is enough to show that

$$(n - m_n + 1/2)(\log n - \log(n - m_n)) \geq m_n \log c + \frac{1}{2} \log m_n,$$

which holds for $c < 1$ and n large enough (since the left hand side is positive and the right hand side negative). This proves (11) for (33).

To finish the proof assume that both

$$\begin{aligned} \max_{\pi \in S_{m_n, n}} \sum_{i=1}^{m_n} Y_{i, \pi(i)} &\leq m_n \log n + \frac{m_n \log \log n}{\lambda} \quad \text{and} \\ Z_{n,k} &\leq \mathbb{E}(Z_{n,k})^\gamma, \quad \text{for each } 1 \leq k \leq \log n + \frac{\log \log n}{\lambda}, \end{aligned}$$

hold for some $\lambda > 1$, which is true for n large enough almost surely. The same calculations as in the proof of the upper bounds in Theorem 1 and (34) yield

$$\text{per } A \leq \left(\frac{n}{m_n} \right)^\gamma n^{(\beta-\gamma)m_n} m_n^{m_n \gamma} (\log n)^{\frac{\beta-\gamma}{\lambda} m_n} e^{m_n \gamma} \left(\log n + \frac{1}{\lambda} \log \log n \right)^{\gamma m_n}$$

and

$$\frac{\log \text{per } A}{m_n \log n} \leq \gamma \frac{\log \left(\frac{n}{m_n} \right)}{m_n \log n} + \beta - \gamma + \gamma \frac{\log m_n}{\log n} + \frac{\beta - \gamma}{\lambda} \frac{\log \log n}{\log n} + \frac{\gamma}{\log n} + \gamma \frac{\log(\log + \frac{1}{\lambda} \log \log n)}{\log n}.$$

The last three terms vanish in the limit and so it remains to prove

$$\limsup_{n \rightarrow \infty} \left(\frac{\log \left(\frac{n}{m_n} \right)}{m_n \log n} + \frac{\log m_n}{\log n} \right) \leq 1.$$

After applying (21) we are left with

$$\limsup_n \left(\frac{n}{m_n} - \left(\frac{n}{m_n} - 1 \right) \frac{\log(n - m_n)}{\log n} \right) \leq 1.$$

which follows from (23). For $\beta \leq 1$ one can repeat the calculations, or simply refer to stochastic domination. \square

The proof of Theorem 14 is based on the following equivalent of Lemma 12.

Lemma 15. *Let ξ be a positive random variable and let A_n be a sequence of $m_n \times n$ matrices whose elements are independent and identically distributed as ξ and which height m_n satisfies (29). For any $0 < \alpha < 1$ and any $\delta > 0$ there exists $r > 0$ with the following property: Almost surely there exists n_0 such that for any $n \geq n_0$ and any pair of integers (k_1, k_2) satisfying $\alpha m_n \leq k_1 \leq m_n$, $\alpha n \leq k_2 \leq n$, and $k_1 \leq k_2$, any $k_1 \times k_2$ submatrix B of A_n satisfies $\text{per } B \geq r^{k_1 k_2^{(1-\delta)k_1}}$.*

Sketch of the proof of Lemma 15. One follows the proof of Lemma 12. In the definitions one needs to write n for the width of the matrix and m_n for the height, for example \mathcal{C}_n is defined as the event that for some pair of integers (k_1, k_2) satisfying $1 \leq k_1 \leq m_n$, $1 \leq k_2 \leq n$ and $k_1 + k_2 \geq \alpha n$ some $k_1 \times k_2$ submatrix of A_n contains only zeros, and \mathcal{B}_n is defined as before. On $\mathcal{B}_n^c \cap \mathcal{C}_n^c$ one has

$$\text{per } B \geq \begin{cases} q^{k_1 \lfloor (1-\delta)k_2 \rfloor!}, & \text{for } k_1 \geq (1-\delta)k_2 \\ q^{k_1 \frac{\lfloor (1-\delta)k_2 \rfloor!}{(\lfloor (1-\delta)k_2 \rfloor - k_1)!}}, & \text{for } k_1 < (1-\delta)k_2. \end{cases}$$

In the first case the logarithm of the right hand side will be bounded from below by

$$k_1 \log q + (1-\delta)k_2 \log((1-\delta)k_2) - (1-\delta)k_2 \geq k_1(\log q - 1 + \log(1-\delta)) + (1-\delta)k_1 \log k_2,$$

and in the second

$$\begin{aligned} k_1 \log q + ((1-\delta)k_2 + 1/2) \log((1-\delta)k_2) - ((1-\delta)k_2 - k_1 + 1/2) \log((1-\delta)k_2 - k_1) + k_1 \\ \geq k_1 \log((1-\delta)k_2) + k_1(1 + \log q), \end{aligned}$$

which is sufficient in both cases. Probability of the event \mathcal{B}_n is estimated as before. For \mathcal{C}_n one can follow the arguments in (22) starting with

$$\mathbb{P}(\mathfrak{C}_n) \leq 2 \sum_{\alpha n/2 \leq k_2 \leq n} \binom{n}{k_2} \sum_{1 \leq k_1 \leq m_n} \binom{m_n}{k_1} \eta^{k_1 k_2}.$$

This inequality follows from the simple fact that $m_n \leq n$ and $k_1 \leq k_2$ imply that $\binom{m_n}{k_2} \binom{n}{k_1} \leq \binom{m_n}{k_1} \binom{n}{k_2}$. \square

Sketch of the proof of Theorem 14. Lower bounds follow from Lemma 15. For the upper bounds again use Markov's inequality and reduce to the case when elements of A_n are parameter 1 Pareto distributed. Similarly as before observe that for any $k_1 \times k_2$ submatrix B , any term in the sum defining $\text{per } B$ can be expanded in $\binom{n-k_1}{m_n-k_1} (m_n-k_1)!$ ways to a term in the sum defining $\text{per } A_n$. Since all elements of A_n are greater or equal than 1 we have

$$\begin{aligned} \log \text{per } B &\leq \log \text{per } A_n - \log \left(\binom{n-k_1}{m_n-k_1} (m_n-k_1)! \right) \\ &\leq \log \text{per } A_n - (n-k_1+1/2) \log(n-k_1) + (n-m_n+1/2) \log(n-m_n) + (m_n-k_1). \end{aligned}$$

Terms $m_n - k_1$, $\log(n - m_n)$ and $\log(n - k_1)$ are $o(k_1 \log k_2)$. After disregarding them, by the proven upper bounds in Theorem 13, almost surely for any $\epsilon > 0$ one has for n large enough

$$\frac{\log \text{per } B}{k_1 \log k_2} \leq \frac{\log n}{\log k_2} \left(\epsilon \frac{m_n}{k_1} + \frac{m_n \log n + (n - m_n) \log(n - m_n) - (n - k_1) \log(n - k_1)}{k_1 \log n} \right).$$

We are left to prove that

$$(35) \quad -(n - k_1) \log(n - k_1) + (n - m_n) \log(n - m_n) + (m_n - k_1) \log n \leq o(m_n \log n),$$

uniformly in $\alpha m_n \leq k_1 \leq m_n$. Since the left hand side (as a function of k_1) is increasing on $(0, n(1 - 1/e))$ and decreasing on $(n(1 - 1/e), 1)$ it suffices to prove (35) for $k_1 = m_n$ when $m_n \leq n(1 - 1/e)$ and for $k_1 = n(1 - 1/e)$ when $n(1 - 1/e) < m_n \leq n$ (assuming $\alpha < 1 - 1/e$ which is of no loss of generality). For $k_1 = m_n$ the left hand side is 0 so there is nothing to prove, and for $k_1 = n(1 - 1/e)$ the left hand side is equal to

$$\frac{n}{e} - (n - m_n)(\log n - \log(n - m_n)),$$

which is not larger than an $o(n \log n)$ term, and this suffices in the case $n(1 - 1/e) < m_n \leq n$. \square

Sketch of the proof of the lower bounds in Theorem 13. The lower bounds for $\beta \leq 1$ case follow directly from Theorem 14 i). For the case $\beta > 1$ one can follow the arguments almost verbatim. In (24) one needs to replace ρn by ρm_n in the upper limit in the sum and in the factor in the end. Moreover the greedy algorithm is run for $k = \rho m_n$ steps, where ρ is defined as in (26) with n replaced by m_n (except under logs). One defines the submatrices B and B^c as before and arrives at the analog of (28)

$$\mathbb{P}\left(\frac{\log \text{per } B^c}{\beta m_n \log n} \leq t\right) \leq \sum_{i=1}^k \mathbb{P}\left(\frac{Q_i}{\log(n - i + 1)} \leq t'\right) \leq \rho m_n \exp(-(1 - \rho)^{1-t'(1+\epsilon)} n^{1-t'(1+\epsilon)}).$$

To finish the proof apply Lemma 15 on B identically as Lemma 12 in Section 4. \square

This example shows that Theorem 1 in general fails when the limit in (2) does not exist. Actually this is possible at arbitrary small oscillations of the sequence in (2). We present the argument for square matrices.

Example 16. Let $S = \{k_i\}$ be a set of positive integers labeled so that $k_{i+1} > 2k_i$. Fix $C_2 > C_1 > \lambda > 1$ and for every $k \geq 1$ define the following sequences of positive real numbers

$$t_k = \exp(\lambda^k), \quad \tilde{p}'_k = \exp(-\lambda^k/C_1), \quad \text{and } p'_k = \begin{cases} \exp(-\lambda^k/C_1), & \text{for } k \notin S \\ \exp(-\lambda^k/C_2), & \text{for } k \in S. \end{cases}$$

Clearly both series $\sum_k p'_k$ and $\sum_k \tilde{p}'_k$ converge, so we can normalize the sequences with the its sums Z and \tilde{Z} respectively and obtain sequences $p_k = p'_k/Z$ and $\tilde{p}_k = \tilde{p}'_k/\tilde{Z}$. Let ξ and $\tilde{\xi}$ be random variables supported on the set $\{t_k\}$ with distributions $\mathbb{P}(\xi = t_k) = p_k$ and $\mathbb{P}(\tilde{\xi} = t_k) = \tilde{p}_k$. Observing that the mappings $t \mapsto \mathbb{P}(\xi \geq t)$ and $t \mapsto \mathbb{P}(\tilde{\xi} \geq t)$ are constant on $(t_k, t_{k+1}]$ and that $\mathbb{P}(\xi \geq t_k) \leq 2p_k$, for all $k \in S$ large enough and for infinitely many $k \notin S$ as well, and $\mathbb{P}(\tilde{\xi} \geq t_k) \leq 2\tilde{p}_k$, for all k large enough, it is easy to see that

$$(36) \quad \liminf_{t \rightarrow \infty} \frac{\log \mathbb{P}(\xi \geq t)}{\log t} = -\frac{1}{C_1/\lambda} < -\frac{1}{C_2} = \limsup_{t \rightarrow \infty} \frac{\log \mathbb{P}(\xi \geq t)}{\log t},$$

$$(37) \quad \liminf_{t \rightarrow \infty} \frac{\log \mathbb{P}(\tilde{\xi} \geq t)}{\log t} = -\frac{1}{C_1/\lambda} < -\frac{1}{C_1} = \limsup_{t \rightarrow \infty} \frac{\log \mathbb{P}(\tilde{\xi} \geq t)}{\log t}.$$

As usual let (A_n) denote a sequence of $n \times n$ matrices on a common probability space with independent elements distributes as ξ . We will prove that

$$(38) \quad \liminf_n \frac{\log \text{per } A_n}{n \log n} \leq C_1, \quad \text{and } \limsup_n \frac{\log \text{per } A_n}{n \log n} = C_2$$

To get the upper bound on \liminf take a sequence (ℓ_i) of positive integers such that $k_i < \ell_i < k_{i+1}$ and that sequences $(\ell_i - k_i)$ and $(k_{i+1} - \ell_i)$ are strictly increasing. Define

integers $n_i = \exp(\lambda^{\ell_i}/C_1)$. By a simple union bound the probability that A_n contains an element t_ℓ , for some $\ell \geq k_{i+1}$ is bounded from above by

$$2n_i^2 p_{k_{i+1}} = \frac{1}{Z} \exp(2\lambda^{\ell_i}/C_1) 2 \exp(-\lambda^{k_{i+1}}/C_2),$$

for i large enough. This expression is summable in i , so almost surely A_{n_i} does not contain elements greater than $t_{k_{i+1}}$ for i large enough. Thus to prove the first inequality in (38) one can assume that elements in A_{n_i} are distributed as $\xi \mathbf{1}_{(\xi < t_{k_{i+1}})}$. Next observe that $\xi \mathbf{1}_{(\xi < t_{k_{i+1}})}$ is stochastically dominated by $t_{k_i} \tilde{\xi}$, that is

$$\mathbb{P}(t \leq \xi < t_{k_{i+1}}) \leq \mathbb{P}(t_{k_i} \tilde{\xi} \geq t).$$

While the inequality is trivial for $t \geq t_{k_{i+1}}$ and for $t \leq t_{k_i}$, for $t_{k_i} < t < t_{k_{i+1}}$ it follows from the fact that for $J \subset (t_{k_i}, t_{k_{i+1}})$

$$\mathbb{P}(\xi \in J) = \frac{\sum_{k: t_k \in J} p'_k}{Z} \leq \frac{\sum_{k: t_k \in J} \tilde{p}'_k}{\tilde{Z}} = \mathbb{P}(\tilde{\xi} \in J),$$

since $p'_k = \tilde{p}'_k$, for $t_k \in J$ and $\tilde{Z} \leq Z$. Thus if \tilde{A}_n is the sequence of $n \times n$ matrices whose elements are identical and distributed as $\tilde{\xi}$ then

$$\liminf_n \frac{\log \text{per } A_n}{n \log n} \leq \lim_i \frac{n_i \log t_{k_i} + \log \text{per } \tilde{A}_{n_i}}{n_i \log n_i} \leq \lim_i \frac{\lambda^{k_i}}{\lambda^{\ell_i}/C_1} + C_1 = C_1.$$

Here the second inequality follows from the upper bounds in Theorem 1 and (37).

To prove the second relation in (38) fix $k \in S$, $\epsilon > 0$ and define the integer $n = n_k = \exp((1 + \epsilon)\lambda^k/C_2)$. We proceed with a greedy algorithm analogous to the one in the proof of the lower bound in Theorem 1. With the probability $1 - (1 - p_k)^n$ there is an element in the first row of $A^{(0)} = A_n$ equal to t_k . On this event take the first such element, remove the corresponding column and the first row from A_n and obtain the $(n - 1) \times (n - 1)$ matrix $A^{(1)}$ which is independent of the first row and is distributed as A_{n-1} . Now repeat the step with $A^{(1)}$ instead of $A^{(0)}$ and proceed recursively as long as one is successful at each step. For $0 < \rho < 1$ one will not be able to proceed till step ρn with probability at most

$$\sum_{(1-\rho)n \leq i \leq n} (1 - p_k)^i \leq \frac{(1 - p_k)^{(1-\rho)n}}{p_k} \leq 3Z \exp\left(-\frac{1-\rho}{Z} \exp(\epsilon\lambda^k/C_2) + \lambda^k/C_2\right),$$

if k is chosen large enough. The right hand side is clearly summable in k and thus almost surely for k large enough and $n = n_k$ constructed as above, the above algorithm will be successful for ρn steps. In that case one can get lower bound on $\text{per } A_n$ as in the proof of the lower bounds in Theorem 1: The matrix at the intersection of the first ρn rows and the removed columns is bounded from below by the product of extracted elements, that is $t_k^{\rho n}$ and the matrix at the intersection of the last $(1 - \rho)n$ rows and non-removed columns is bounded from below by $((1 - \rho)n)!$ (since all of its elements are greater or equal than 1). Therefore for k large enough and n constructed as above $\text{per } A_n \geq t_k^{\rho n} ((1 - \rho)n)!$. Since $\log((1 - \rho)n)!/(n \log n) \rightarrow 1 - \rho$ and

$$\frac{\log t_k^{\rho n}}{n \log n} \geq \frac{\rho \log t_k}{\log n} \geq \frac{\rho \lambda^k}{(1 + \epsilon)\lambda^k/C_2} \rightarrow \frac{C_2 \rho}{1 + \epsilon},$$

we obtain that almost surely

$$\limsup_{n \rightarrow \infty} \frac{\log \text{per } A_n}{n \log n} \geq \frac{C_2 \rho}{1 + \epsilon} + 1 - \rho.$$

By sending $\rho \rightarrow 1$ and $\epsilon \rightarrow 0$ we get C_2 as the lower bound on the lim sup, and by Theorem 1 and (36) it is equal to C_2 .

Acknowledgments The author would like to thank Professor Sourav Chatterjee for suggesting this problem and for generous help and guidance with this paper.

REFERENCES

- [1] A. J. van Es and R. Helmers. Elementary symmetric polynomials of increasing order. *Probab. Theory Related Fields*, 80(1):21–35, 1988.
- [2] V. L. Girko. Inequalities for a random determinant and a random permanent. *Teor. Veroyatnost. i Mat. Statist. Vyp.*, 4:48–57, 1971.
- [3] V. L. Girko. A sharpening of certain theorems for a random determinant and permanent. *Teor. Veroyatnost. i Mat. Statist.*, (7):28–32, 1972.
- [4] Marshall Hall, Jr. Distinct representatives of subsets. *Bull. Amer. Math. Soc.*, 54:922–926, 1948.
- [5] P. Hall. On representatives of subsets. *Journal of the London Mathematical Society*, s1-10(1):26–30, 1935.
- [6] Svante Janson. The numbers of spanning trees, Hamilton cycles and perfect matchings in a random graph. *Combin. Probab. Comput.*, 3(1):97–126, 1994.
- [7] V. S. Korolyuk and Yu. V. Borovskikh. Random permanents and symmetric statistics. In *Probability theory and mathematical statistics (Kiev, 1991)*, pages 176–187. World Sci. Publ., River Edge, NJ, 1992.
- [8] H. B. Mann and H. J. Ryser. Systems of distinct representatives. *Amer. Math. Monthly*, 60:397–401, 1953.
- [9] Henryk Minc. *Permanents*, volume 9999 of *Encyclopedia of Mathematics and its Applications*. Addison-Wesley Publishing Co., Reading, Mass., 1978. With a foreword by Marvin Marcus, *Encyclopedia of Mathematics and its Applications*, Vol. 6.
- [10] Grzegorz Rempala and Jacek Wołowski. Limiting behavior of random permanents. *Statist. Probab. Lett.*, 45(2):149–158, 1999.
- [11] Grzegorz A. Rempala and Jacek Wołowski. Strong laws of large numbers for random permanents. *Probab. Math. Statist.*, 22(2, Acta Univ. Wratislav. No. 2470):201–209, 2002.
- [12] Grzegorz A. Rempala and Jacek Wołowski. *Symmetric functionals on random matrices and random matchings problems*, volume 147 of *The IMA Volumes in Mathematics and its Applications*. Springer, New York, 2008. With a foreword by Douglas N. Arnold and Arnd Scheel.
- [13] T. Tao and V. Vu. On the permanent of random bernoulli matrices. 2008. <http://arxiv.org/abs/0804.2362>.

MATHEMATICS DEPARTMENT, UNIVERSITY OF CALIFORNIA, BERKELEY, CA 94720, USA

E-mail address: tantun@math.berkeley.edu